Data and Pattern Mining of Federal Computer Science Research Investments

Project Title	Data and Pattern Mining of Federal Computer Science Research Investments
Project Summary	Explore patterns and changes in sponsored scientific research in selected technical disciplines within computer science based upon open-source databases of research projects across the federal government.
Country	United States

Project Description

NSF's Cyber Physical Systems (CPS) program funds research in topics where computers touch the physical world, including transportation, agriculture, manufacturing, homes, etc. Three important open sources data for this project are the Vanderbilt CPS-VO, a repository of project data on cyber-physical systems, and the National Science Foundation Awards Database and archive of program solicitations. These databases have largely not been the object of scientific investigation for what the data may tell about a history of scientific and technological advances and innovations from investments by the National Science Foundation. The ability to use tools for data mining, data analytics, and visualization is important as is the ability to write clearly and produce meaningful infographics. The intern will work with researchers at the Vanderbilt University's Institute for Software Integrated Systems, who manage and curate the CPS-VO.

Required Skills or Interests

Skill(s)
Analytical writing
Data visualization

Additional Information

Other relevant skills include Clustering tools (e.g., Carrot Search Lingo4G), Data visualization, Drupal, Machine learning, Network science, Text analytics (e.g., Apache SOLR).

Students at Vanderbilt University are particularly encouraged to apply.

Language Requirements

None